View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by YorkSpace

# THE EFFECTS OF COMMON FATE GROUPING ON SUPRATHRESHOLD DEPTH ESTIMATES

**MICHAEL MARIANOVSKI**

**A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE**

**GRADUATE PROGRAM IN BIOLOGY**
**YORK UNIVERSITY**
**TORONTO, ONTARIO**

**JUNE 2016**

**© MICHAEL MARIANOVSKI, 2016**

## ABSTRACT

When two vertical lines are perceived to form the boundaries of a common object, observers underestimate their separation in depth. This disruption in perceived depth depends directly on the perceived grouping via closure of the resultant figure. Here I tested if this effect generalises to motion based grouping cues, specifically common fate.

Stimuli were presented on a mirror stereoscope and an on screen ruler was used to record depth magnitude estimates. It was found that depth estimates were accurate over a range of suprathreshold disparities, for *both* static and moving stimuli. Attempts to strengthen the grouping cues had no impact on depth magnitude estimates. This was true even when a more complex biological motion stimulus was used. These results show that previously reported reductions in perceived depth are not generalizable to grouping via common motion. It appears that this phenomenon only occurs when the spatial layout suggests a common object.

# TABLE OF CONTENTS

# LIST OF FIGURES

# INTRODUCTION

## Stereopsis

Stereoscopic vision refers to the perception of depth based on information derived from the fact that the two eyes are laterally offset in the head. This results in two slightly different images of the same scene projected onto each retina. This slight difference is referred to as binocular disparity and it is used by the visual system to compute 3-D structure (Howard & Rogers, 2012; Wheatstone, 1838). Binocular disparity is typically present when viewing scenes in the natural environment, but can also be created by presenting separate, offset images to each eye individually. This technique is widely used in cinema, entertainment technology as well as psychophysical experimentation.

Very high resolution depth information can be extracted from binocular disparity. For example, stereoacuity (depth discrimination thresholds), for practiced observers can be as low as 2 to 6 arcsecs. Thresholds are somewhat higher (and more variable) for the general population but remain precise — approximately 80% of the population are able to discriminate disparities of 30 arcsecs or more (Coutant & Westheimer, 1993). Several low-level stimulus attributes have been shown to influence these discrimination thresholds including, luminance, contrast, size, location, and duration (Berry, Riggs, & Duncan, 1950; Christophers & Rogers, 1994; Ogle & Weil, 1958; Woo & Sillanpaa, 1979). Such properties are often referred to as early or low-level properties because they are known to be processes at relatively early stages in the visual cortex (i.e. V1).

Stereoacuity tests were performed as early as 1900 using the three rod test designed by Helmholtz (Helmholtz, 1925). In later studies, more akin to modern psychophysical experiments,

observers viewed two pairs of rods either sequentially or simultaneously and chose the pair with a depth offset (Fan, Brown, & Yap, 1996). Modern stereoacuity tests such as the RANDOT™ test (Stereo Optical Co.), which contains a series of random dot stereograms, can be performed quickly and efficiently but lack the fine experimental control necessary for psychophysical experiments (Fawcett & Birch, 2000).

In addition providing high resolution depth discrimination information, stereopsis can also provide information about the *amount* of depth between two stimulus elements. Depth estimates made in this way are relatively precise over a wide range of binocular disparities (Ogle, 1952a, 1952b). However, they are not always veridical (based on predicted stereoscopic geometry) (Bülthoff, Fahle, & Wegmann, 1991; Foley & Richards, 1972; Ogle, 1953). As mentioned above, many studies have evaluated the effects of stimulus attributes on stereoacuity, but it is not clear that these results can be extrapolated to depth magnitude estimation, which is a suprathreshold judgment (Foley, Applebaum, & Richards, 1975). In fact, several studies have shown that stereoacuity and suprathreshold depth estimates are not always affected the same way by the same stimulus attributes. For example, Schor and Howarth (1986) showed that while disparity thresholds decrease with spatial frequency below 2.5 cpd, suprathreshold depth estimates remain robust down to 0.5 cpd. Similarly, Patel et al. (2009) showed that simple stimulus manipulations which elevate disparity thresholds do not always disrupt depth magnitude estimates. To date, there is no complete explanation to account for these relationships between threshold and suprathreshold stereoscopic processing.

**Effects of configuration**

As described above low level stimulus attributes such as luminance, contrast, size, location, and duration can alter stereoscopic thresholds (Berry et al., 1950; Christophers & Rogers, 1994; Ogle & Weil, 1958; Woo & Sillanpaa, 1979). There is also evidence that more complex stimulus attributes, specifically configuration, can influence stereoacuity. The first evidence of this was reported by Werner, 1937 who while studying the effect of practice on observers' ability to use a stereoscope, found that connecting two vertical lines with horizontal lines made it very difficult to detect depth between the vertical lines. Later, McKee (1983) demonstrated that stereothresholds were elevated when two target elements at different disparities were connected by straight lines. Her experiments showed that thresholds were elevated greater than 6-fold when two vertical lines were connected by horizontal lines to create a rectangular figure. Additionally, it has been shown that various spatial connections between the vertical lines also result in degraded stereoacuity. These can be explicit (for instance, an H configuration) (McKee, 1983; Mitchison & Westheimer, 1984) or implicit as exemplified by a configuration resembling two square brackets ([ ]) (Mitchison & Westheimer, 1984) . McKee (1983) argued that configuration-based depth degradation is a consequence of conflict between disparity detection and binocular fusion. She proposed that information from binocular disparity detectors is pooled when elements are connected. Consequently, two vertical lines on different depth planes may be assigned a single disparity value, which is averaged across the rectangular figure they create (see also Fahle & Westheimer, 1988). A series of experiments by Mitchson and Westheimer (1984) systematically assessed the impact of configuration on threshold elevation and found that physically connecting target elements was not necessary to disrupt depth percepts. They proposed that when two isolated lines offset in depth are viewed, the visual

system uses a hypothetical fronto-parallel plane as a reference from which to make relative disparity judgments. When the lines are then connected or just merely made coplanar (with supporting reference cues) the reference plane is redefined to be parallel with the target elements. This configuration places the target elements equidistant from the reference plane and thus no relative depth between them is detected (see Glennerster & McKee, 1999). The influence of a reference plane was also demonstrated by Fahle & Westheimer (1988) who investigated how depth detection thresholds between two dots are affected by the addition of intervening dots. They found that intervening dots significantly elevated thresholds (compared to isolated dots) and as more dots were added, increasing the salience of a reference plane, thresholds were elevated further. However, if the intervening dots were moved off the imaginary plane created by the outermost dots, threshold elevation was not as large. Fahle & Westheimer (1988) proposed that two mechanisms influence stereoscopic depth detection in these stimuli. The first is a local mechanism which is capable of processing only disparity information between two target elements. The second is a more global mechanism which is influenced by global stimulus features such as configuration and other elements in the scene. Another explanation for the observed stereoacuity degradation was provided by Zalevski et al. (2007) who argued that the configuration-based elevated thresholds can be attributed to cue conflict with perspective. That is, in the stimuli used by McKee (1983) and others described above, the vertical line height was held constant, however, in natural circumstances lines further from the observer should subtend a smaller visual angle than those closer (perspective foreshortening). Zalevski et al. (2007) reported that when the length of the vertical lines was adjusted to account for perspective, thresholds improved. However, the disruptive effect of the connecting lines was not completely eliminated (see also Lu, Tjan, & Liu, 2006).

In a recent paper, Deas and Wilcox (2014) argued that the explanations for the interaction between figural processing and stereopsis described above are incomplete because they only consider low-level stimulus properties and local disparity interactions. Instead, higher level perceptual interpretations of the stimulus (specifically based on perceptual grouping) must be involved. Further, while there was some reference to perceptual grouping in the previous literature, no attempt was made to explicitly manipulate or quantify the extent of this grouping. In their studies, Deas and Wilcox (2014) used stimuli similar to those used by McKee (1983) and others, but employed a suprathreshold, depth magnitude estimation task to evaluate the impact of perceptual grouping on depth perception. More specifically, in their first series of experiments they investigated the effect of perceptual closure on depth perception. Closure is a well-studied Gestalt organizational principle originally defined by Koffa (1935) as the transformation of lines, into surfaces bounded by lines, when the lines are connected. Their results showed that when two vertical lines are connected to form a closed rectangle, suprathreshold depth estimates are reduced, echoing McKee's (1983) threshold results. In experiments where Deas and Wilcox (2014) directly manipulated perceived closure, their results showed that depth estimates were inversely related to the appearance of the stimulus as a unified object. This provided an alternative explanation for the configural depth-degrading effects previously proposed by Mitchson and Westheimer (1984), Fahle & Westheimer (1988) and Zalevski et al. (2007) who focused on low level disparity processing. Deas and Wilcox (2014) argued, rather, that their results are evidence of a top-down influence on disparity processing—higher level object based interpretations of the stimulus promote disparity smoothing (or pooling) and result in degraded depth estimates.

**Gestalt grouping and depth degradation**

Deas and Wilcox (2014) showed that depth magnitude estimates are reduced as a consequence of perceptual closure, a Gestalt grouping principle. This potentially implies that other Gestalt grouping cues may also lead to configuration based depth degradation. Consequently, I hypothesized that any manipulation that enhances the perceptual grouping of two elements should reduce the percept of depth between them. Therefore, the aim of this thesis was to investigate whether object based depth degradation is a general consequence of Gestalt grouping.

The focus of Deas & Wilcox's (2014; 2015) studies was closure; the perceptual grouping of elements that form a closed contour even if the contours are partially occluded or incomplete (Koffa, 1935). Early research on the perceptual impact of closure assumed that it was an all-or-nothing quality. For instance in the case of a rectangular object, the vertical lines are either connected (closed) or not. However, more recent research has shown that perceived closure falls on a continuum, which corresponds to the degree of spatial support. For example, a rectangular figure segmented by a gap along one side will be less likely to be perceived as a closed figure as the gap width is increased (Gillam, 1975). Regardless of how closure is defined, some spatial support is required to form a 2-D surface from 1-D contours (Elder & Zucker, 1993, 1994; Koffa, 1935). Thus, the presence of explicit or implied connectivity is an important contributor to the percept of objectness from closure (see Figure 1) (Hadad & Kimchi, 2008; Kovacs & Julesz, 1993; Shevelev, Kamenkovich, & Sharaev, 2003). Like closure, several other Gestalt grouping principles rely on spatial support (contours or edges), these include common region, continuity, proximity, similarity of shape or orientation and symmetry (Wagemans et al., 2012). In contrast, as described below, the Gestalt cue of common fate facilitates grouping elements

based on similar temporal coincidence and relies more heavily on changes in features over time rather than space. If the reduced depth estimates reported by Deas and Wilcox (2014) were a result of perceptual grouping then the same reduced depth estimates should be observed as a consequence of common fate grouping. If, however, their results were a consequence of another feature of closure (such as spatial support) then common fate grouping may not lead to the same reductions in perceived depth. The latter result would indicate that depth degradation is not a general consequence of Gestalt grouping cues.

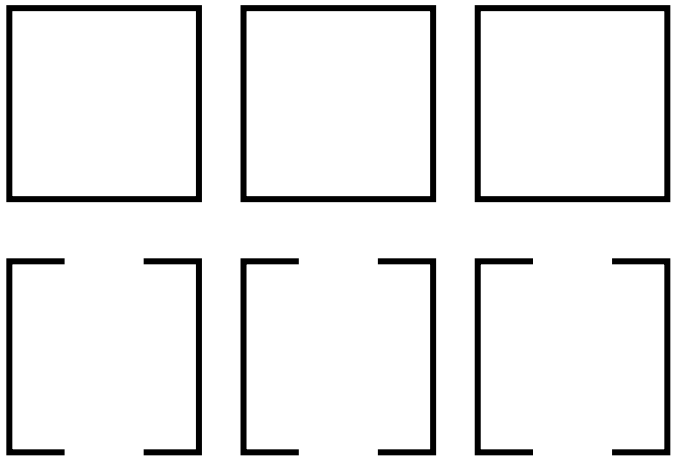

*Figure 1. Illustrations of explicit (top row) and implied (bottom row) closure. Note that even without being physically connected to each other, the vertical line elements of the pattern in the bottom row are perceptually grouped to form the outer edges of three rectangles.*

**Common fate**

Common fate is a well-known perceptual grouping principle based on motion. According to this principle, elements moving in the same direction with a similar velocity are perceived as belonging to the same entity (Wertheimer, 1938). This powerful organizational principle can be readily observed in everyday life. For instance, in the collective movement of a flock of birds. Common fate is also believed to be a special case of similarity grouping based on similar temporal changes in feature value (such as brightness or location) rather than absolute feature

value (Sekuler & Bennett, 2001). The perceptual impact of common fate is very robust and can be achieved at a large range of relative figure/ground speeds, direction and orientation of movement (Sekuler, 1990; Wist, Ehrenstein, & Schraus, 1998).

When first defined, common fate was considered one of the most important organizational principles used to segregate a figure from its surroundings (Wertheimer, 1923). Subsequently researchers have consistently demonstrated that observers can readily segment figure from ground by using common fate as a grouping cue. (Braddick, 1974; Sekuler, 1990; Wist et al., 1998). For instance, Braddick (1974) showed that when a section of random dots was moved against a background of random dots, the moving section was perceived in front of the background. Relevantly to this thesis, the moving dots were perceived as a coherent surface. This indicates that common fate grouping can potentially lead to figural interpretations of stimuli which is a requirement for object based disparity pooling. Another demonstration of this effect involved moving an object behind a series of occluding strips. When the object was static it was very difficult to recognize but when in motion it clearly appeared to be a coherent object/surface (Burr & Ross, 1986; Sperling, Landy, Dosher, & Perkins, 1989). Experiments like these show that common fate grouping facilitates perceptual organisation of elements into objects. Consequently, it may lead to reduced depth estimates by object based disparity pooling as proposed by Deas and Wilcox (2014).

In this thesis I built on the results of Deas and Wilcox (2014) who argue that top-down influences of shape perception lead to degraded depth percepts. I used stimuli designed to create coherent objects via common fate, to investigate whether object based depth degradation is a wide-spread effect of Gestalt grouping principles.

# GENERAL METHODS

## Observers

A total of thirty-five observers were tested in the experiments described here. Thirty were inexperienced (naïve) observers who were recruited through the York University Undergraduate Research Pool (URPP), five observers had extensive experience in depth magnitude estimation tasks. All participants had normal or corrected-to-normal visual acuity and stereoacuity of at least 40 seconds of arc as tested using the RANDOT™ stereoacuity test. Observers were also assessed using a depth magnitude estimation task with a stimulus configuration similar to the one used in Experiments 1-4 (outlined below). Observers who were not able to scale their depth estimates with increasing disparity were excluded from participating. Prior to participating, the interpupillary distance (IPD) of all observers was measured using a ruler. The research protocol used here and in subsequent experiments was approved by the York University research ethics board.

## Apparatus

Stimuli were generated using the Psychtoolbox package for MATLAB® on a Microsoft® Windows® XP computer. They were presented on a pair of CRT monitors using a mirror stereoscope with a viewing distance of 50 cm. The monitor resolution was 1600 x 1200 pixels with a refresh rate of 100Hz. Depth estimates were made using a Logitech® gamepad to adjust the length of a virtual ruler (horizontal line) on the screen. This method has been previously validated by Hartle & Wilcox (2014) for use with similar stimuli.

## Theoretical depth from disparity

Using conventional binocular geometry, disparity was converted to theoretical depth to facilitate easier comparison with reported depth estimates. A formula was used which relates

disparity to predicated depth: Predicted Depth = ($d * D^2 / IPD$), where d is the relative disparity, D is the viewing distance and IPD is the interpupillary distance (Howard & Rogers, 2012). The average IPD of the observers was used for the calculation.

## EXPERIMENT 1

A number of studies have shown that depth information is degraded as a result of stimulus configurations in which vertical lines are connected by horizontal ones (Deas & Wilcox, 2014; McKee, 1983; Mitchison & Westheimer, 1984). As described in the Introduction section, perceptual closure has been shown to modulate this depth-degradation; as perceived closure increases, depth estimates become smaller (Deas & Wilcox, 2014). The aim of Experiment 1 was to investigate whether reduced depth estimates are a general consequence of Gestalt grouping cues, more specifically, if they also apply to motion-based grouping. To evaluate this, suprathreshold depth estimates were obtained for stimuli under common motion and compared to those made for static stimuli. To facilitate comparison with previous experiments, a modified version of the stimuli used by McKee (1983) and later adapted by Deas and Wilcox (2014) was used.

### Observers

Ten observers participated in Experiment 1. Eight were naive stereoscopic observers. Two were experienced observers.

### Stimuli

The stimuli comprised four white (68.8 cd/$m^2$) vertical lines (3.4° x 0.1°) positioned at the centre of a grey display (10.9 cd/$m^2$). The horizontal separation between the lines was 1.9°. A 0.2° wide white horizontal bar (virtual ruler) was positioned 6° below the four vertical lines. The length of the bar was randomized at the beginning of each trial. Six disparities were tested (0, 0.08, 0.17, 0.25, 0.34 and 0.42°). They were selected by taking into account the viewing distance and display resolution to create predicted depths of approximately 0, 0.6, 1.2, 1.9, 2.5,

and 3.1 cm. The two outer lines remained in the fixation plane (zero disparity) in all conditions. Four conditions were tested by manipulating the motion of the two central lines (Figure 2).

A) Vertical motion: The two central lines oscillated vertically about the midpoint of the display.
B) Horizontal motion: The two central lines oscillated horizontally about the midpoint of the display.
C) Motion in depth: The two central lines oscillated in depth (along the z-axis). The oscillation ranged from the initial disparity of each line to a preselected crossed disparity.
D) Static lines: All four lines remained stationary throughout a trial.

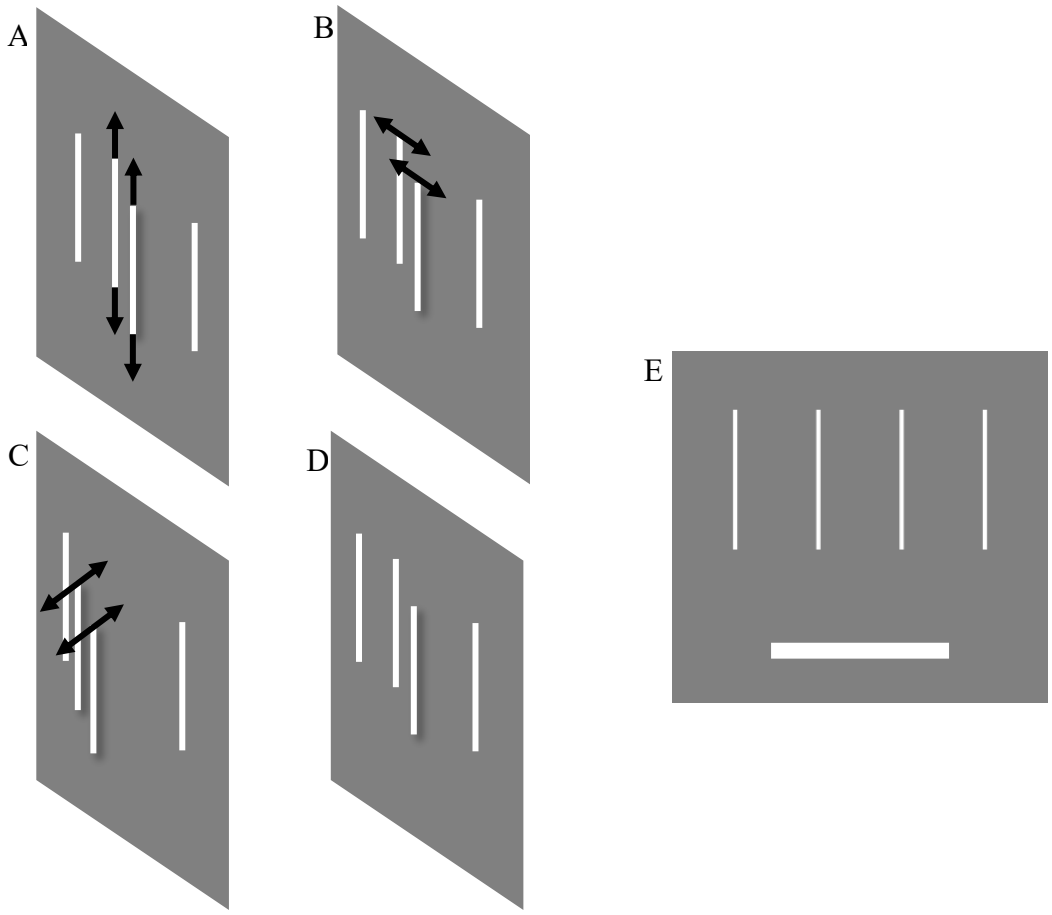


*Figure 2. Perspective representation of stimuli used in Experiment 1. Each panel represents a different condition: (A) Vertical motion, (B) Horizontal motion, (C) Motion in depth and (D) Static lines. The black arrows indicate the oscillatory motion path. Observers judged the depth between the two central lines in each condition. (E) Illustrates one eye's view of the stimulus. The horizontal white bar below the vertical lines was used as a virtual ruler to make relative depth magnitude judgements.*

**Procedure**

Each trial began with the presentation of a white fixation cross (0.7 x 0.7°) with zero disparity for 700 ms followed by the stimulus. On a given trial one of the two central lines, randomly selected on each trial, was presented at one of the six test disparities. In all conditions observers were asked to estimate the relative depth between the two central lines by using two

buttons on a gamepad to increase or decrease the length of the virtual ruler. The stimulus remained on the screen from the onset of the trial until the observer recorded their response by pressing a third button. Each condition-disparity combination was repeated 10 times and divided into two sessions (5 repeats per session) for a total of 240 trials.

**Results and Discussion**

Figure 3 shows observers' depth estimates as a function of predicted depth. As disparity increases there is a corresponding monotonic increase in estimated depth for all four conditions. In all cases there is a small overestimation at depths below 2 cm; at larger disparities estimates are more variable but are still tightly grouped within 0.5 cm of each other. A repeated measures ANOVA was used to analyse the results. A Mauchly's test of sphericity showed that the assumption of sphericity was violated for the main effect of condition and disparity as well as the interaction of condition x disparity, $p<0.0001$ so the Greenhouse-Geisser correction was applied. There was a significant main effect of disparity ($F(1.38,12.47) = 115.22$, $p<0.0001$; $\eta^2 = 0.93$) and of condition ($F(1.14,12.73) = 4.75$, $p=0.038$; $\eta^2 = 0.34$). There was also a significant condition x disparity interaction ($F(4.78,43.09) = 3.42$, $p<0.0001$; $\eta^2 = 0.27$). Simple main effects of condition (with a Bonferroni adjustment) revealed that the interaction was driven by significant differences between the motion in depth condition and the static lines condition at the two largest disparities ($p=0.005$ and $p<0.0001$, respectively). No other individual comparisons approached significance.

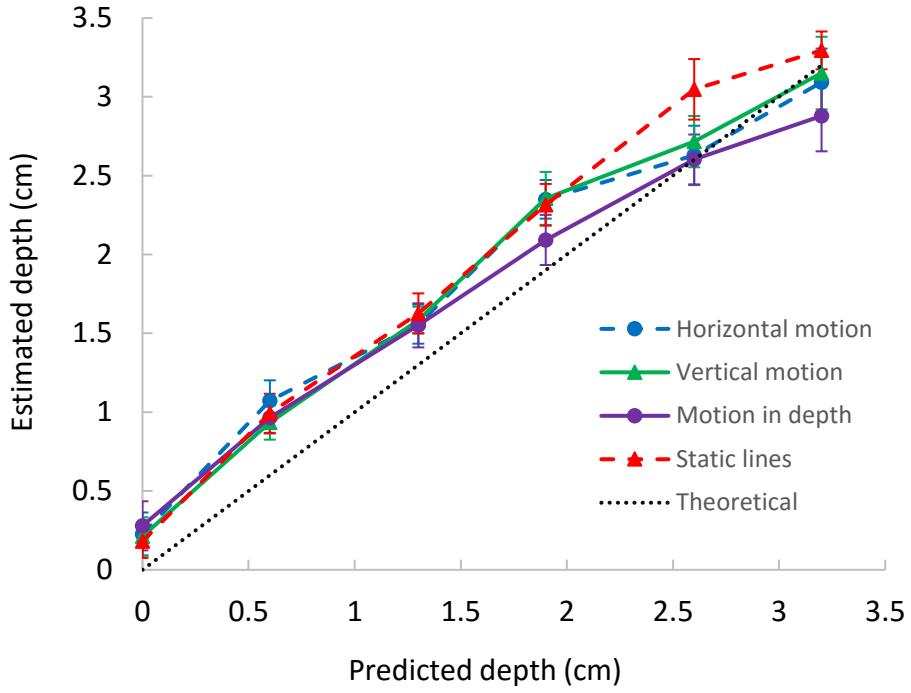


*Figure 3. Estimated depth is plotted as a function of predicted depth for four motion conditions in Experiment 1: horizontal (blue), vertical (green), in-depth (purple) and static (red). The dotted black line represents the theoretically predicted depth. Error bars show +/- 1 SEM.*

It was predicted that by introducing motion to the central lines they would appear grouped and perceptually form the boundaries of a single object. As argued by Deas and Wilcox (2014), interpretations of a stimulus as an object can result in object based disparity pooling and reduced depth estimates. Consequently, I hypothesized that stimuli grouped by motion would be subject to a reduction in perceived depth. If this hypothesis was correct then the results should have shown relatively higher depth estimates for the static lines condition and relatively lower depth estimates for the three motion conditions. However, observers reported very similar amounts of depth in the four conditions; only the motion in-depth condition had lower estimates than the static lines condition. Two questions are raised by these results: why were depth

estimates for all four conditions so similar? And why was the motion in depth condition the only one to show reduced depth estimates?

One potential answer to the first question is that the common motion present in the stimuli was not sufficient to perceptually group the lines to the same extent as closure. Consequently, the stimuli were not interpreted as coherent objects. If this was the case then it would not be surprising that the depth estimates were similar across conditions. However, depth estimates were lower for the motion in depth condition at the two largest disparities. Care was taken to control for any differences between the four conditions used in this experiment so that they differed only in the motion of the central lines. However, the motion in depth condition inherently has stimulus features that differ from the other motion conditions which may explain this pattern of results: First, (i) the disparity pedestal between the fixation plane and reference target line (the central line farthest from the observer) changes as the lines move closer and farther from the observer. In comparison, there is no disparity pedestal in the other three conditions (since the reference line is always on the fixation plane). Stereothresholds have been shown to be elevated as a function of depth pedestal (Blakemore, 1970; Ogle, 1953). Therefore, it is possible that the presence of a changing disparity pedestal disrupted depth perception and contributed to reduced depth estimates. Additionally, (ii) the height of the two central lines was not adjusted as they were repositioned in depth. In the natural environment, it is well known that for an image of constant physical size, retinal image size is proportional to distance from the observer (Howard & Rogers, 2012); as an object becomes closer, its retinal image gets larger. Since the length of the lines was not adjusted as a function of their motion, their retinal size remained the same. This may have created an illusion of changing size due to size constancy. That is, if two objects have the same retinal size, but one appears to lie closer to the observer in

depth, then it will appear smaller (Howard & Rogers, 2012). Further, perceptual constancy also occurs in the disparity domain (depth constancy). That is, the retinal disparity between two physical objects with a constant 3-D separation, should increase as they move closer to an observer (Wallach & Zuckerman, 1963). In this experiment, the disparity between the central lines was kept constant throughout the motion. Consequently, the central lines may have appeared to move closer to each other in depth as they moved closer to the observer. Note that under this illusion created by depth constancy the relative depth between the two central lines would only appear veridical at the moment that the reference line is on the fixation plane. During the rest of the motion path (the majority of a trial) the relative depth would appear reduced. It follows that the estimated separation in depth of the lines may have been reduced as a consequence. In summary, it is likely that the presence of these two (i-ii) factors would add noise to the stimulus and may contribute to the reduced depth estimates observed in the motion in depth condition in Experiment 1.

# EXPERIMENT 2

## Introduction

I hypothesized that the common motion manipulation in Experiment 1 was not sufficiently salient to cause perceptual grouping of the target lines. If so, adding segregation cues may increase the extent to which they are perceived as a single object. In Experiments 2 and 3 I systematically increased the salience of the common motion and segregation of the two target lines by adding texture and manipulating the contrast polarity of the pattern.

In Experiment 2 a texture of randomly scattered dots was added to the fixation plane. A small portion of the dots was randomly positioned on the slanted plane created by the two central lines when they were offset in depth. This texture created a strong fronto-parallel reference plane at 0° of disparity as well as a slanted interpolated surface between the central lines of the stimulus (see Saidpour, Braunstein, & Hoffman, 1994). These texture cues provided some spatial support for figural interpretations of the stimulus. Note that because the texture did not fall on the contours which would connect the endpoints of the vertical lines, they should not have introduced closure. In Experiment 3 the colour of the lines as well as the size of the texture elements was manipulated in a further attempt to increase perceptual grouping of the two central lines.

## Observers

Ten observers participated in Experiment 2. Eight were naive stereoscopic observers. Two were experienced observers who also participated in Experiment 1.

**Stimuli**

As in Experiment 1, the stimuli comprised four white vertical lines (3.4° x 0.1°) positioned at the centre of a grey display. Small dots (0.02° in diameter) were added to the fixation plane at 0° of disparity to create a textured reference plane (Figure 4A). They filled a rectangular area spanning the entire display horizontally and 5.2° above and below the vertical lines (as to not interfere with the virtual ruler). A subset of the dots (1%) appeared on the plane created between the two central lines (Figure 4B). Restrictions were placed on the location of the dots to avoid the formation of illusory contours above and below the lines. The same four conditions described in Experiment 1 were tested: horizontal motion, vertical motion, motion in depth and static lines. The procedure described in Experiment 1 was used.

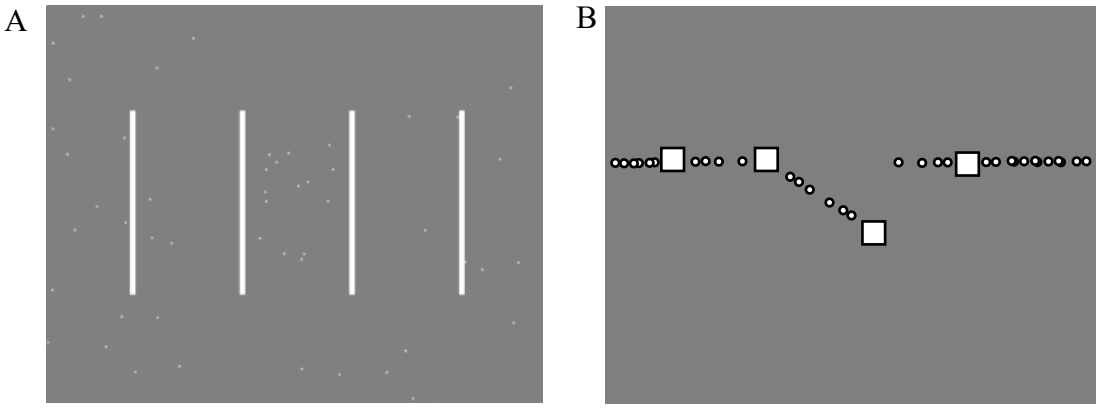


*Figure 4. (A) Illustration of one eye's view of the stimulus used in Experiment 2. (B) Top view of stimulus used in Experiment 2. Note, the top view is not drawn to scale.*

**Results**

Figure 5 shows observers' depth estimates as a function of predicted depth. As disparity increases there is a corresponding increase in depth estimates. All four conditions show a small overestimation at every disparity level but estimates remain very similar to theoretically predicted values. A repeated measures ANOVA was used to analyse the results. A Mauchly's test

of sphericity showed that the assumption of sphericity was violated for the main effect of condition and disparity as well as the interaction of condition x disparity, $p<0.0001$ so the Greenhouse-Geisser correction was applied. There was a significant effect of disparity ($F(1.55,14.02) = 233.58$, $p<0.0001$; $\eta^2 = 0.96$). , with no significant main effect of condition (($F(2.05,18.45) = 3.09$, $p=0.068$; $\eta^2 = 0.25$) or condition x disparity interaction ($F(4.57,41.19) = 2.66$, $p=0.073$; $\eta^2 = 0.23$).

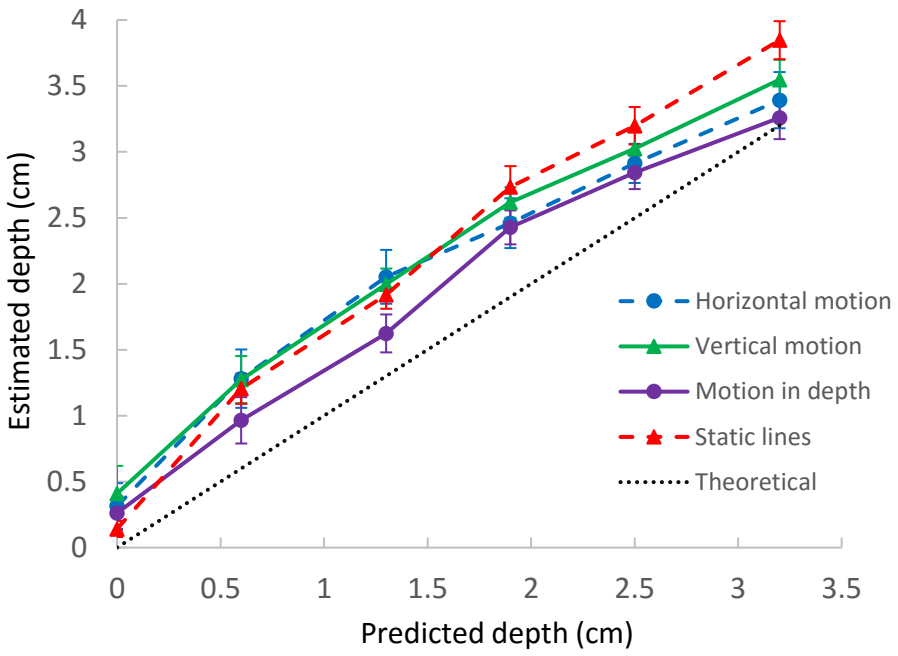


*Figure 5. Estimated depth is plotted as a function of predicted depth for four motion conditions: horizontal (blue), vertical (green), in-depth (purple) and static (red). The dotted black line represents the theoretically predicted depth. Error bars show +/- 1 SEM.*

# EXPERIMENT 3

## Observers

Ten observers participated in Experiment 3. Seven were naive stereoscopic observers one of whom also participated in Experiment 2. Three were experienced observers one of whom also participated in Experiment 1.

## Stimuli

The stimuli consisted of a modified version of the stimuli used in Experiment 2. The size of the dots was increased from a diameter of 0.02° to 0.07°. The colour of the dots positioned between the target lines was changed from white to black (0 cd/$m^2$) while the remaining dots remained white. The colour of the two central lines was also changed from white to black while the two outer lines remained white (Figure 6). The same four conditions tested in Experiment 1 were assessed: horizontal motion, vertical motion, motion in depth and static lines. The procedure described in Experiment 1 was used.

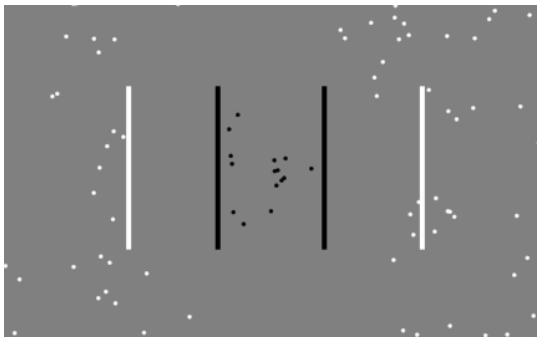

*Figure 6. Illustration of one eye's view of the stimulus used in Experiment 3.*

## Results

Figure 7 shows observers' depth estimates as a function of predicted depth. As disparity increases there is a corresponding increase in estimated depth in all four conditions. The

horizontal and vertical motion condition as well as the static lines condition show a slight overestimation. These three conditions also show very similar depth estimates. The motion in depth condition shows relatively accurate depth estimates up to 2 cm of depth. Above this point depth is underestimated. Notably, compared to the other conditions, depth estimates are reduced for the motion in depth condition at all non-zero disparity values. A repeated measures ANOVA was used to analyse the results. A Mauchly's test of sphericity showed that the assumption of sphericity was violated for the main effect of condition and disparity as well as the interaction of condition x disparity, $p<0.0001$ so the Greenhouse-Geisser correction was applied. The ANOVA revealed a significant effect of disparity ($F(1.69,15.21) = 173.29$, $p<0.0001$; $\eta^2 = 0.95$) and condition ($F(2.10,18.90) = 18.43$, $p<0.0001$; $\eta^2 = 0.67$). There was also a significant interaction between condition and disparity ($F(5.44,49.04) = 10.13$, $p<0.0001$; $\eta2 = 0.52$) therefore simple main effects were run (a Bonferroni correction was applied). The comparisons revealed that the motion in depth condition differed significantly from the horizontal motion, vertical motion and static lines conditions ($p=0.01$, $p=0.002$ and $p= 0.002$, respectively). The remaining conditions were not statistically different from one another.

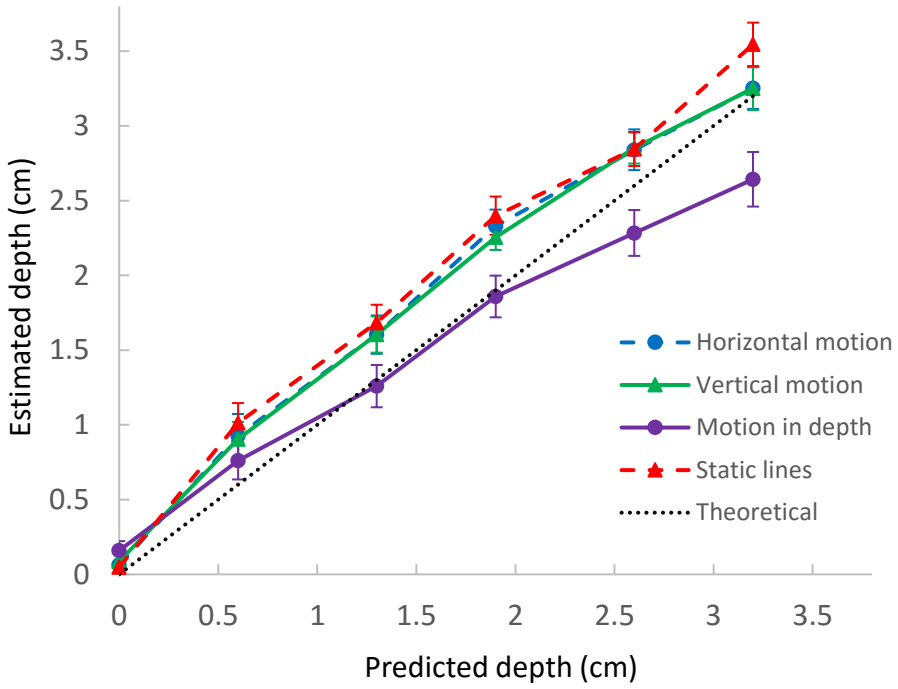


*Figure 7. Estimated depth is plotted as a function of predicted depth for four motion conditions: horizontal (blue), vertical (green), in-depth (purple) and static (red). The dotted black line represents the theoretically predicted depth. Error bars show +/- 1 SEM.*

**Discussion**

I predicted that the texture-defined fixation plane and luminance contrast would serve to enhance grouping of the central line pair in Experiment 3. In turn, this would decrease depth estimates between the lines as a consequence of disparity pooling between them. In spite of the presence of this additional segregation information, depth estimates in the moving and static conditions were similar (with exception of the motion depth in condition discussed below). These data show that the results of Experiments 1 and 2 were not due to the strength of the segmentation of the lines. Instead, it appears that common motion does not induce the same reductions in depth estimates as does closure (perhaps due to the lack of spatial support which may be necessary of object based disparity pooling).

The results of Experiment 3 show a significant reduction in depth estimates in the motion in depth condition, unlike Experiment 1, in this study it occurred at all non-zero test disparities. As outlined in the Results and Discussion section of Experiment 1 this drop in estimated depth may be caused by stimulus features which are inherent to the motion in depth conditions: changing disparity pedestal, perspective cue conflict and depth constancy. However, the underestimation of depth for the motion in depth condition was larger and more consistent in Experiment 3 than in Experiments 1 and 2. It likely that this occurred due to the presence of a well-defined reference plane and high contrast between the target lines and the reference lines. The presence of these two factors may increase the strength of the disparity signal by reducing correspondence ambiguity. That is, the difference in contrast polarity (black on grey vs white on grey) may help the stereoscopic system correctly match the line features in the two eye's views. In turn, the depth constancy illusion described in the Results and Discussion section of Experiment 1 may appear more pronounced which can account for the more consistent lower depth estimates (see Cormack, Stevenson, & Schor, 1991). An alternative explanation is based on 'good disparity continuation' defined by Deas and Wilcox (2015) as the grouping of elements which lie on a smooth disparity gradient. They showed that when disparity varies smoothly in a linear formation of dots through depth, depth estimates between elements at each end of the series of dots is reduced. It is possible that the texture elements present in Experiment 3 (specifically the dots on the slanted plane between the two central lines) created a smooth disparity gradient between the two target lines which enhanced grouping and therefore reduced depth estimates. While more consistent depth reduction was observed in this experiment relative to Experiments 1 and 2, it is important to note that it cannot be attributed to common fate motion alone, for it only occurs in the more complex motion in depth conditions. Further, as outlined in

the Discussion section of Experiment 4, the size of the reduced depth effect in this study remains much weaker than that created via closure.

## EXPERIMENT 4

In designing the preceding experiments I assumed that minor alterations to the line stimuli (e.g. dimensions, spacing) would not impact the putative grouping effects. In light of the absence of the degraded depth effect in Experiments 1-3, it was important to replicate the disruptive effect of closure on perceived depth reported by Deas and Wilcox (2014) using the stimuli employed here. In addition, by replicating their experiments with the stimuli used here, I would be able to directly compare the impact of grouping by closure and by common fate on depth from disparity.

### Observers

Ten observers participated in Experiment 4. Nine were naive stereoscopic observers one of whom also participated in Experiment 2. One was an experienced observers.

### Stimuli

Four conditions were tested using a modified version of the stimulus presented in Experiment 1 (Figure 8):

A) Static lines: All four lines remained stationary throughout the duration of a trial.
B) Static rectangle: The two central lines were connected by two horizontal lines to form a rectangular figure slanted in depth.
C) Vertical motion (lines): The two central lines oscillated vertically about the midpoint of the display.
D) Vertical motion (rectangle): The slanted rectangle created by connecting the two vertical lines, oscillated vertically about the midpoint of the display.

The methodology described in Experiment 1 was used here.

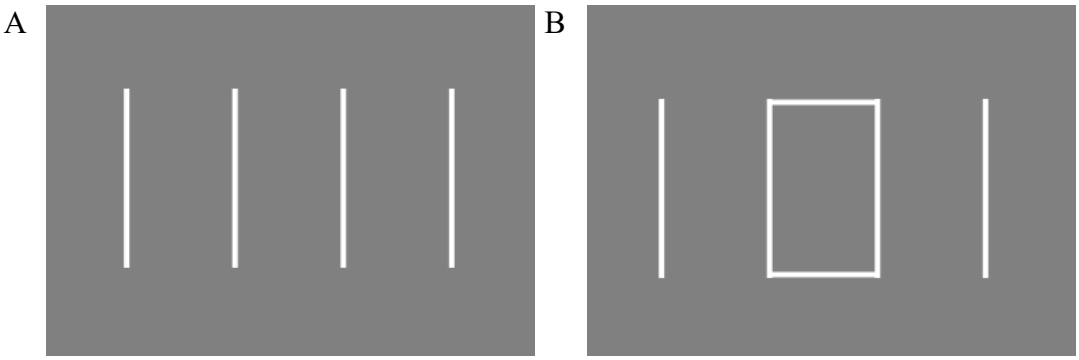


*Figure 8. Illustration of one eye's view of the stimulus used in Experiment 4. (A) Represents two conditions, the central lines either oscillated vertically or were static. (B) Represents two conditions, the central rectangle either oscillated vertically or was static.*

**Results and Discussion**

Figure 9 shows observers' depth estimates as a function of predicted depth. Depth estimates increased monotonically for all four conditions as a function of disparity. The two isolated line conditions show a slight overestimation of depth while the two rectangle conditions show depth underestimation compared to theoretical predictions. Most notably, there was a clear and systemic reduction in depth estimates in the two rectangle conditions. A repeated measures ANOVA was used to analyse the results. A Mauchly's test of sphericity showed that the assumption of sphericity was violated for the main effect of condition and disparity as well as the interaction of condition x disparity, $p<0.0001$ so the Greenhouse-Geisser correction was applied. There was a main effect of disparity ($F(1.36,12.30) = 60.04$, $p<0.0001$; $\eta^2 = 0.87$) and condition ($F(1.13,10.19) = 19.51$, $p<0.0001$; $\eta^2 = 0.68$). The ANOVA also revealed a significant condition x disparity interaction ($F(2.64,23.75) =11.74$, $p<0.0001$; $\eta2 = 0.56$) therefore simple main effects were run (a Bonferroni correction was applied). All comparisons between the two lines and two rectangle conditions were significant: static lines and static rectangle ($p=0.005$), static lines and moving rectangle ($p=0.009$), moving lines and static rectangle ($p=0.001$), and moving lines and

moving rectangle ($p=0.010$)). There was no significant difference within the two rectangle conditions or the two line conditions.

The reduction in depth estimates for the two rectangular figure conditions, regardless of motion, clearly consolidates the results of Experiments 1-3. Motion based grouping, even with additional cues such as salient reference planes and contrast similarity, does not lead to reduced depth estimates. However when closure is introduced by explicitly connecting the lines, the reduction in depth percepts is evident. This shows that the absence of the degraded depth effect in the previous experiments is not due to trivial changes in stimulus configuration.

As noted in Experiment 3, there was a significant reduction in depth percepts for the motion in depth condition when matching ambiguity was reduced, and texture elements were positioned along the slanted surface between the target lines. I suggested that this was primarily due to disparity-based factors, including disparity grouping. This argument is supported by the results of the current study; for comparison, the motion in depth results of Experiment 3 are re-plotted in Figure 9 along with the results of Experiment 4. It is clear that there is a more substantial reduction in perceived depth for stimuli grouped by closure than by motion in depth. This confirms that the reduced depth estimates observed for that condition are likely due to the factors described above, and not to grouping via common fate.

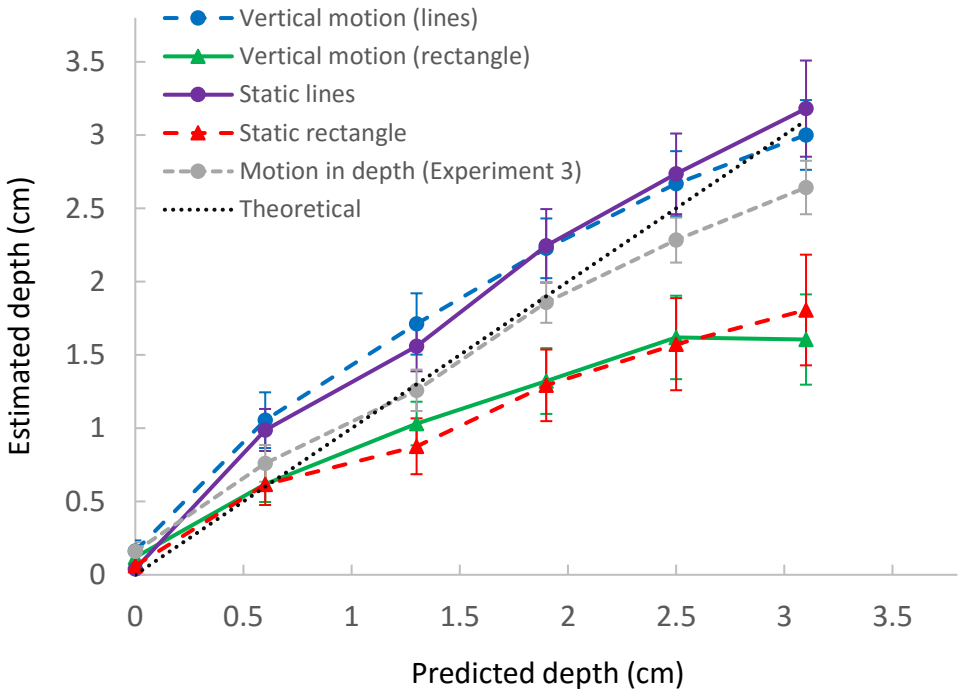


*Figure 9. Estimated depth is plotted as a function of predicted depth for five conditions: vertically oscillating lines (blue), vertically oscillating rectangle (green), static lines (purple), static rectangle (red) and the motion in depth condition form Experiment 3 (grey). The dotted black line represents the theoretically predicted depth. Error bars show +/- 1 SEM.*

## EXPERIMENT 5

Experiments 1-4 provide convincing evidence that common fate grouping does not reduce depth estimates. However, relatively simple line stimuli were used in these experiments. It is possible that a stimulus with more naturalistic motion-based grouping cues may enhance grouping, and therefore disrupt depth percepts. This hypothesis is supported by studies which show that more complex point-light biological stimuli disrupt depth perception (Bulthoff, Bulthoff, & Sinha, 1998; Lu et al., 2006). For example, Bulthoff et al. (1998) showed that observers are more likely to indicate that joints are coplanar (even if they are not) when those joints are on the same limb as opposed to different limbs of point light figures. Also, Lu et al. (2006) showed that when hand-to-elbow distances between two biological walker arms differ in length, participants show elevated discrimination thresholds. However, these studies did not directly assess either stereothresholds or depth magnitude. Consequently, their results may not generalize to the suprathreshold depth estimation tasks I present here. I hypothesized that the introduction of biological motion to point-light biological stimuli, would create strong grouping between joints positioned on the same limb and degrade depth estimates between them.

### Observers

Ten observers participated in Experiment 5. Eight were naive stereoscopic observers three of whom also participated in Experiment 3. Two were experienced observers who also participated in Experiment 2 and 3.

### Stimuli

The stimulus was a point light walker created from 13 black dots representing the head, shoulders, elbows, wrists, hips, knees and ankles (Figure 10) (Hodgins, 2005; van Boxtel & Lu, 2013). The biological figure was presented in a stationary walking motion facing the right side of

the screen. The walker height (head to ankle) was 8.6°. The maximum walker width (ankle to ankle at peak stride) was 2.8°. The joints were divided into three groups and distributed on three depth planes. The left arm and leg joints were offset in depth with an uncrossed disparity of 0.07°. The dot representing the head was fixed at zero disparity and the right arm and leg had a crossed disparity of 0.07°. Two target dots, between which depth estimates were made, were highlighted in red with a smaller black centre. A white vertical bar spanning 0.2° in width was presented 6° to the left of the walker. The length of the bar was randomised at the beginning of each trial. Five disparities were tested (0, 0.11, 0.21, 0.32 and 0.42). They were selected by taking into account the viewing distance and display resolution to create predicted depths of approximately 0, 0.8, 1.5, 2.3 and 3.1 cm. Four conditions were tested:

A) Biological motion: The figure 'walked' in place, facing the right side of the screen. The target dots were the joints of the right knee and ankle.
B) Static point-light figure (PLF): A static biological figure was presented. It consisted of a single frame taken from the point light walker in mid-stride. The target dots were the joints of the right knee and ankle.
C) Scrambled: The dots followed the same path as in condition (A) but their starting position was scrambled in 2-D. Constraints on dot placement were used to assure that the distance between the two target dots was on average the same as that between the knee and ankle in conditions (A) and (B) and that no dots interrupted the space between the two target dots during the motion.
D) Two isolated dots: Two target dots were offset vertically in 2-D in the centre of the screen. This offset was consistent with the distance between the knee and ankle in the previous conditions.

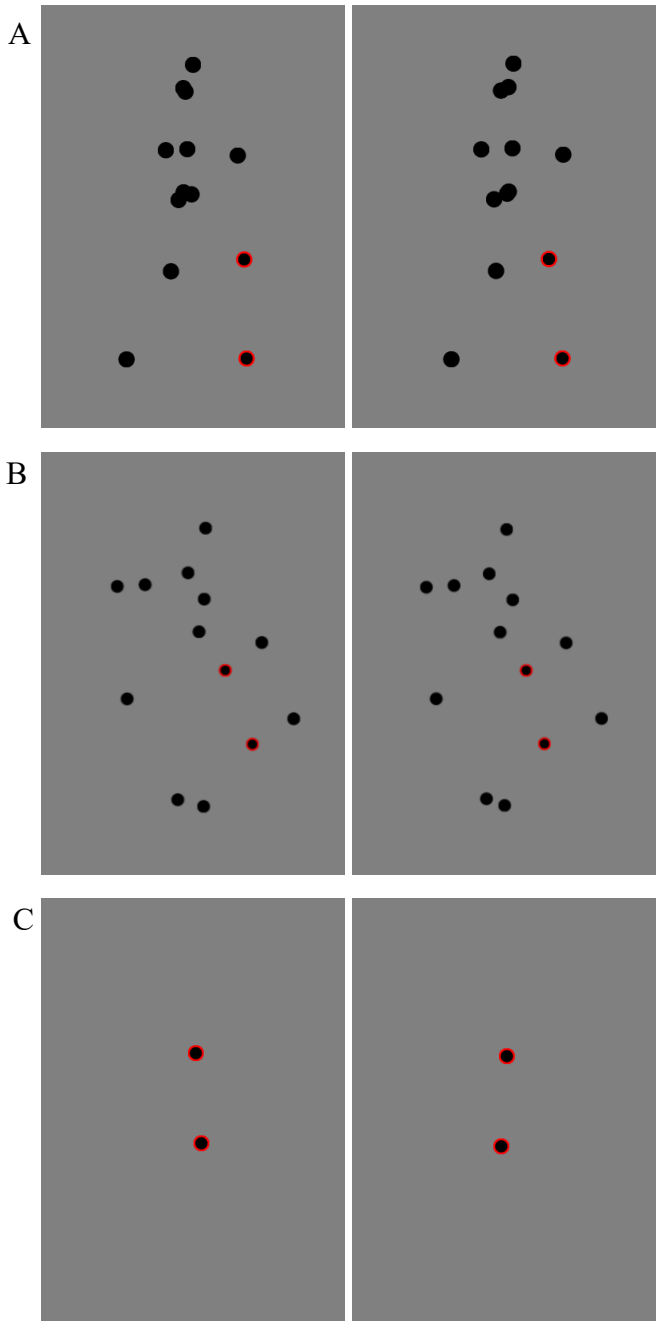


*Figure 10. Stereograms of the stimulus used in Experiment 5 arranged for crossed fusion. The black dots represent joints/head - the two red dots are the target joints. Each panel shows a different condition: (A) Represents two conditions, the dots either followed their joint path (biological motion) or were static as displayed above (static PLF); (B) Scrambled dots and (C) Two isolated dots. Note the size of the dots in this figure was increased for demonstration purposes.*

**Procedure**

Each trial began with the presentation of a white fixation cross (0.7 x 0.7°) for 700 ms followed by the presentation of the stimulus. On a given trial, one of the test disparities was randomly chosen and added to the initial crossed disparity of one of the target dots. This resulted in a separation in depth between the two target dots. The stimulus remained on the screen until the participant indicated how much depth they saw between the two target dots by adjusting the length of the virtual ruler using the gamepad. Each of the four conditions was presented at each of the five disparities and repeated five times per session. A total of 200 trials were conducted equally divided between two sessions.

**Results and Discussion**

Figure 11 shows observers' depth estimates as a function of predicted depth. As disparity increases there is a corresponding monotonic increase in perceived depth for all four conditions. While it appears that the most accurate estimates are obtained in the static PLF condition, all four conditions show depth estimates relatively close to theoretical predictions. A repeated measures ANOVA was used to analyse the results. A Mauchly's test of sphericity showed that the assumption of sphericity was violated for the main effect of condition and disparity as well as the interaction of condition x disparity, $p<0.0001$ so the Greenhouse-Geisser correction was applied. The ANOVA results showed a main effect of disparity ($F(1.18,10.63) = 97.07$, $p<0.0001$; $\eta^2 = 0.91$) with no effect of condition ($F(1.20,10.83) = 3.23$, $p<0.095$; $\eta^2 = 0.26$) or interaction between condition and disparity ($F(3.70,33.37) = 1.32$, $p=0.21$; $\eta^2 = 0.13$).

It was predicted that depth estimates in the biological figure conditions would be reduced; especially in the biological motion condition which had compelling biological and motion based grouping cues. That is, recognition of the stimulus as a biological figure was

expected to disrupt processing of conflicting disparity information. Estimates for the scrambled and isolated dots condition were expected to be veridical due to a lack of grouping cues. Instead, the results show that even under naturalistic motion and with biological form cues, observers are able to ignore any potential grouping cues and extract accurate disparity information from the stimulus.

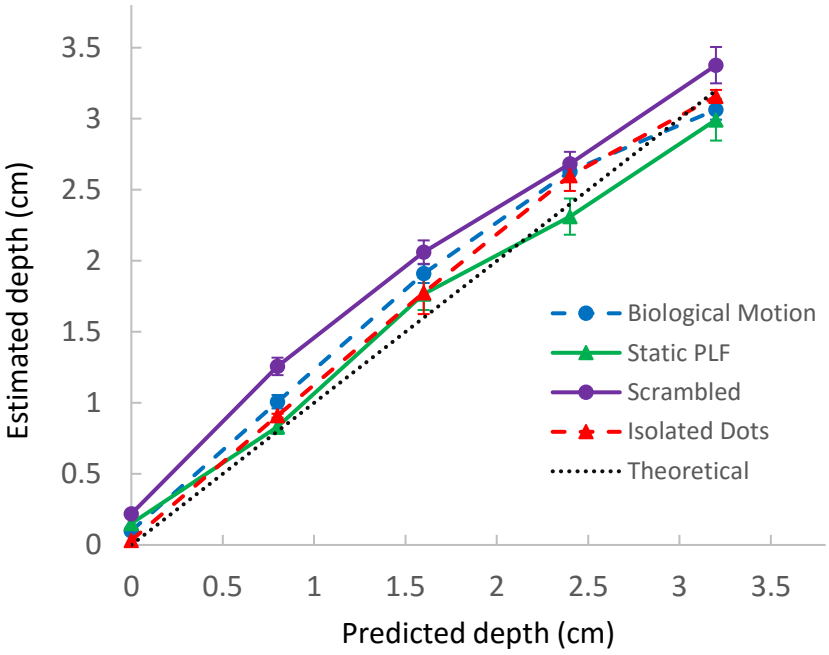


*Figure 11. Estimated depth is plotted as a function of predicted depth for four conditions: biological motion (blue), static point-light figure (green), scrambled dots (purple) and static dots (red). The dotted black line represents the theoretically predicted depth. Error bars show +/- 1 SEM.*

# GENERAL DISCUSSION

## Summary

The experiments described here assess if the disruptive effects of perceptual organization on perceived depth are a general phenomenon, common to grouping cues other than closure. This work was motivated by recent research by Deas and Wilcox (2014) who proposed that the perceptual interpretation of a stimulus as an object, based on Gestalt grouping cues, impacts perceived depth, reducing depth magnitude estimates. Specifically, their results showed that the depth-degrading configural effects were modulated by perceptual closure (Deas & Wilcox, 2014). That is, the salience of figural grouping via perceptual closure was inversely related to perceived depth.

I hypothesised that if the reported disruption in perceived depth is a general outcome of perceptual organization then it should also occur when elements are perceptually organised via other grouping cues. I reasoned that if depth-degradation resulted from a grouping cue other than closure, specifically one that significantly differed from closure, then this would be a first step towards showing that the effect is a general outcome of perceptual grouping. The grouping cue, common fate, was selected for this purpose as it relies on spatiotemporal coincidence unlike closure which relies on spatial support.

In Experiment 1 participants estimated the depth between two vertical lines. Common motion of the two lines was used to promote perceptual grouping. The results showed that the amount of depth perceived between these lines was equivalent to that perceived when viewing static lines. In Experiments 2 and 3 the addition of reference plane elements, and manipulation of luminance contrast, served to help segregate the central line pair and provide a more stable frame of reference for binocular disparity. It was shown that this did not have an effect. That is, depth

estimates for static lines were the same as those for their moving counterparts except when the elements moved in depth. Particularly in Experiment 3, depth estimates for the motion in depth condition were reduced. This was attributed to disparity based factors and not common fate. An experiment using simple line stimuli was conducted (Experiment 4) to directly compare the difference between grouping by closure and by common fate on depth estimates. This study replicated the results of Deas and Wilcox (2014) and confirmed that closure reduces perceived depth, however, no such effect was observed due to common fate grouping. A biological motion stimulus was used in Experiment 5 to investigate whether a more naturalistic motion stimulus would be more susceptible to the grouping effects observed by Deas and Wilcox (2014). However, despite multiple motion based grouping cues (biological motion and common fate) depth estimates remained accurate. Taken together, the results of these experiments suggest that common fate grouping does not influence depth estimates in the same way as closure. It also shows that perceptual organisation does not always lead to disruption in depth perception as originally hypothesized.

**Object based grouping**

In 1983 McKee showed that when two vertical lines, at different disparities, are connected by two horizontal lines to form a rectangle, stereoscopic thresholds are elevated. She proposed that this was a consequence of averaging/pooling the disparity signal between the vertical lines to assign a single disparity value to the rectangle as a whole. However, she did not specify the conditions under which disparity pooling would occur. Recently, Deas and Wilcox (2014) argued that disparity pooling occurs specifically due to the perceptual interpretation of the stimulus as an *object* (which they called object based disparity pooling) (Deas & Wilcox, 2014). Conversely, the results of Experiments 1-4 show that no depth degradation occurred as a result of

introducing common motion to the stimuli. I propose that one explanation for the difference between the results reported here, and those of Deas and Wilcox (2014) is that common fate, unlike closure, does not create a quality of 'objectness' in perceptually grouped elements.

To understand why this is the case it is important to differentiate between two types of perceptual grouping which I term 'cluster segmentation' and 'object grouping. Cluster segmentation involves the perceptual grouping of elements in a visual scene which share some common quality (for example, temporal, spatial or physical). Object grouping occurs when a perceptually grouped cluster of elements possesses spatial support that explicitly or implicitly creates closure between them (for instance, in the form of a closed contour). I propose that, in motion-grouped stimuli, without the presence of closure (spatial support), object grouping does not occur. Instead, the moving elements are segmented as a cluster. Therefore, without perceptual interpretation of the stimulus as an object, object based disparity pooling does not occur.

Previous investigations of configural influences on depth perception commonly used stimuli consisting of closed contours, such as a rectangle (Deas & Wilcox, 2014; McKee, 1983; Mitchison & Westheimer, 1984; Zalevski et al., 2007). The key difference between these stimuli and the ones used here is the presence of spatial support which facilitates the figural interpretation of the stimuli (or object grouping). In simple rectangular figures spatial support is usually present in the form of L-junctions (corners). The importance of spatial support in object interpretations has been well documented in literature. Notably, Attneave (1954) argued that along contours, psychologically important information is concentrated in regions with high magnitude of curvature (such as corners) as opposed to being uniformly distributed along the contour. Experiments in which contours are systematically removed have supported this

proposal. For example, recognisability of figures was shown to decrease when corners were masked compared to when edges were masked (Shevelev et al., 2003). In related experiments, Kovacs and Julesz (1993) showed that even without corners, closed contours provide a strong cue for figural interpretations of stimuli. Their stimuli consisted of contours, formed from a series of Gabor patches, embedded in a background of randomly positioned Gabor patches. They found that when the distance between Gabor elements that form a contour was increased, detectability of open contours (a line of elements) was significantly worse than that of closed contours (a circular formation). That is, the shape of Gabor elements forming a closed contour was easily segmented form the background, but when the contour was open it blended into the background. Experiments like these (Attneave, 1954; Feldman & Singh, 2005; Kovacs & Julesz, 1993; Saarinen & Levi, 1999), show that spatial support in the form of both corners and closed contours facilitates object-based grouping. I argue that the absence of spatial support in the stimuli used in Experiments 1-4 make interpretations of the figure as a cohesive object less likely. This in turn discourages object based disparity pooling across the stimulus. Thus, due to the failure to perceive the stimulus as a coherent object, depth estimates for the moving stimuli are equivalent to those made for static stimuli.

Further evidence of the difference between cluster segmentation and object grouping is provided by examples from the structure-from-motion literature. Structure-from-motion refers to the facilitation of 3-D shape recognition via motion of 2-D retinal projections. This is especially evident for partially occluded objects which may be unrecognizable when static but readily identifiable when moving. This motion-based figural interpretation seems to implicate common fate in grouping based object perception. However, I argue that common motion of the object is not solely responsible for its recognition as an object. Consider, for example, the camouflage of a

tiger which functions remarkably well to group the tiger with its surroundings and reduce the visibility of its form. However, when the tiger begins to move it can easily be spotted. The common fate based grouping of the tiger's markings contributes to its detection and recognition, but it is the spatial support created by the tiger's outline (and the associated accretion/deletion cues) which leads to its recognition as a single animal (object). For comparison, a flock of flying birds is not seen as a single object but rather as a cluster of objects because, although they possess common motion, they do not maintain a coherent closed bounding contour. Similarly, Burr and Ross (1986) showed that the 3-D structure of objects moving behind a series of vertical occluders can be determined with high accuracy. I argue that this occurs because the objects possess a contour which facilitates their figural interpretations. This supports my argument that the spatial support (contour) of a moving object facilitates segregation of that object from its surroundings. Therefore, I propose that for moving elements, common fate is necessary but not sufficient for grouping the elements to form a common object.

I have argued that accurate depth estimates can be made between elements undergoing common motion because such stimuli are not subject to object based disparity pooling. This assumes that in my experiments observers did not perceive stimuli grouped by common fate as objects, but did perceive stimuli grouped by closure as objects. To test this assumption, in a follow-up study, eight participants (seven of whom participated in at least one of Experiments 1-4) were shown the twelve stimuli from Experiments 1-3 along with the static rectangular stimulus from Experiment 4. They were asked to rate the stimuli on a scale of 0 to 10, to indicate the extent to which the two target vertical lines appear to be the outer edges of a single rectangular object. The stimuli were presented stereoscopically in random order under the same experimental conditions used in the previous experiments. Ratings were reported by adjusting a

number visible on the screen using a gamepad. Each of the thirteen conditions was presented five times at two disparities (0 and 0.42°: no disparity and high disparity).

Figure 12 shows the average rating for each of the thirteen stimuli tested at the two disparities. The large overlap of the 95% confidence intervals indicates that disparity does not have a significant effect on the extent to which the stimuli appear object-like. Importantly, Figure 12 also shows that while some of the motion conditions appeared more object-like than the static line conditions, the motion based grouping is never rated as high as the static rectangle conditions. This result confirms that, at least for the stimuli used here, grouping by motion does not lead to object based figure-ground segmentation to the same extent as grouping by closure. This difference may be sufficient to discourage object based disparity pooling in the former but facilitate it in the latter.

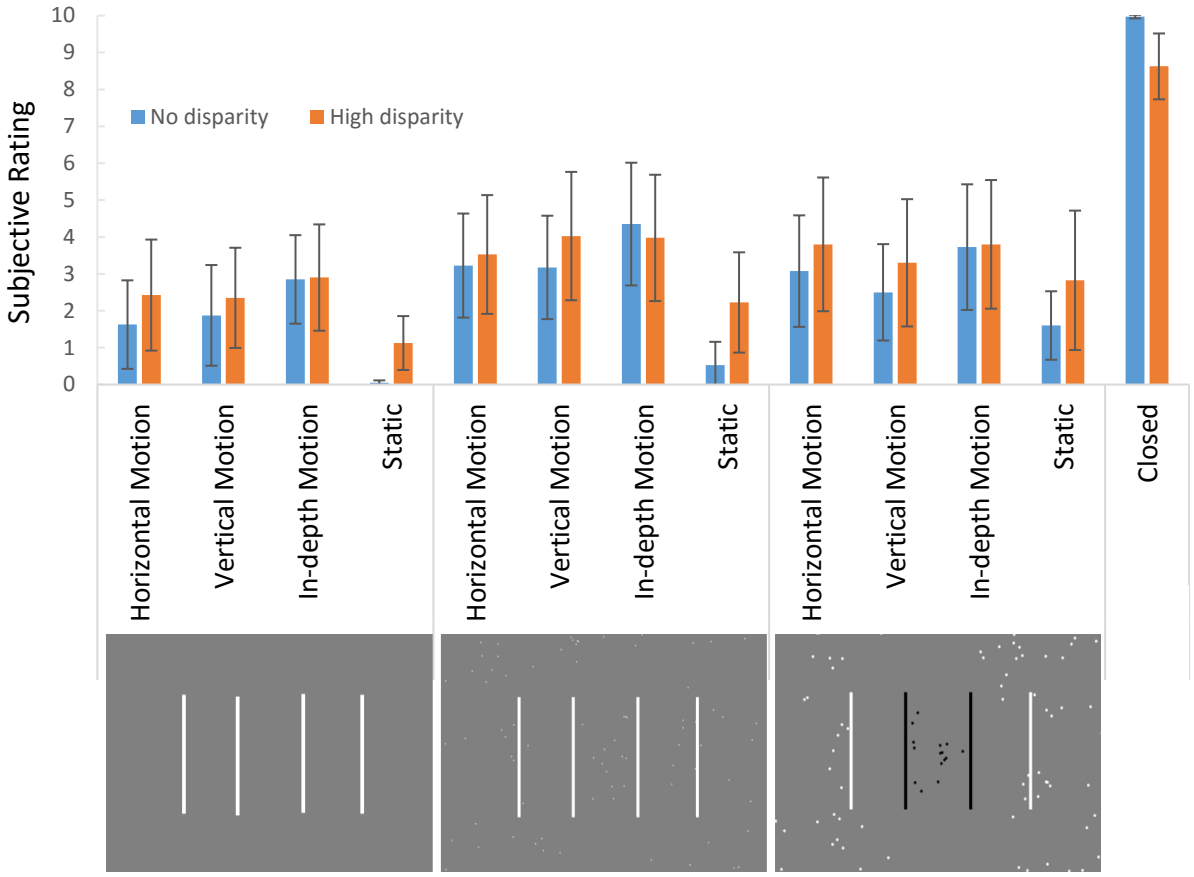


*Figure 12. Average subjective ratings for the thirteen stimulus configurations. Ratings range from 0 (not an object) to 10 (closed object). The stimuli are shown below each rating. Error bars represent 95% confidence intervals.*

**Naturalistic Stimuli**

In Experiment 5 I evaluated observers' depth estimates using biological motion stimuli. The results showed that despite the presence of several potential grouping cues (biological motion and common motion), participants were able to accurately estimate the depth between same-limb joints when they were offset in depth. The biological motion stimulus consisted of dots at strategic joint locations. There was no explicit spatial support present, for instance, in the form of a body shaped contour. Consequently, I argue that the joints were potentially segmented

into a cluster resembling biological form, but the lack of spatial support prevented object grouping and thus object based disparity pooling.

As described in the introduction section of Experiment 5, there is some evidence in the literature that biological motion degrades some aspects of disparity processing (Bulthoff et al., 1998; Lu et al., 2006). These investigators argue that prior expectations of biological form are integrated with sensory information and impede accurate depth estimates due to cue conflict. That is, recognition of the point-light walker as a human figure serves to impose form and meaning onto the stimulus. This sense of form may be so strong that conflicting stereopsis information is perceptually ignored. A possible reason for the apparent discrepancy between these previous results and the results of Experiment 5 are differences between the tasks and stimulus configurations. In the experiments performed by Bulthoff et al. (1998) and Lu et al. (2006), participants were not directly asked to estimate 3-D distances. In such circumstances the observers may not be as sensitive to small disparities. Alternatively, in Experiment 5 when participants were asked to directly estimate the depth between joints, they may have been primed to focus on the disparity present between the joints. This may have led them to consciously ignore cue conflict created by the biological based grouping and precisely process the disparity signal. Interestingly, this priming (to estimate depth) existed in the experiments performed by Deas and Wilcox (2014). As already discussed, their results showed clear reductions in depth estimates despite this priming. This shows that unlike cue conflict, object based disparity pooling cannot always be consciously ignored.

**Ecological considerations**

Since its discovery, common fate has been thought to be applicable to a wide range of conditions which extend beyond grouping elements under common motion (Wertheimer, 1923).

More recently the concept of 'common motion' has been expanded to include temporal coincidence. That is, grouping elements based on common temporal change in feature value (Sekuler & Bennett, 2001). For example, grouping elements by common luminance changes as discussed in the introduction (Sekuler & Bennett, 2001). In such cases, common fate has been redefined as generalized common fate, a special case of similarity grouping based on similar temporal coincidence. Common fate may therefore be important for defining events such as luminance changes across a scene. That is, rather than grouping elements into objects, it may be more useful to group them into a cluster of objects (such as a cluster of objects behind a shadow) (see van den Berg, Kubovy, & Schirillo, 2011). The benefit of grouping elements into a cluster has been demonstrated for similarity grouping. For instance, presenting objects of the same colour to an observer facilitates their enumeration (Halberda, Sires, & Feigenson, 2006). Likewise, estimating the mean size of a cluster of objects can be facilitated by grouping based on colour similarity (Chong & Treisman, 2005). Grouping by common fate may be useful for similar reasons (see Levinthal & Franconeri, 2011).

The importance of grouping elements into a cluster rather than a cohesive object can be appreciated when considering the potential costs of erroneous grouping. Marr and Poggio (1976) proposed a computational model of stereopsis which used smoothing operations as a means to solve the correspondence problem. Such a smoothing constraint is plausible because much of the visual world varies smoothly and only a small proportion is composed of boundaries which are discontinuous in depth. Therefore, for objects which contain relatively little disparity variation relative to their distance to the observer, a smoothing heuristic is quite useful. However, if such a constraint relied on perceptual grouping in order to organise the visual scene into objects, it would be important to avoid false identification of an object as this could lead to loss of valuable

depth information. I have argued here that, common motion of elements is not necessarily an indication that those elements belong to a single object. Instead, it may be more useful to group those elements into a cluster of elements that have some common temporal property. Therefore disparity smoothing over elements grouped by common fate may lead to significant information loss as those elements may themselves be objects with large disparity discontinuities between them.

**Conclusion**

I have described how Gestalt grouping cues affect recovery of 3-D information in simple displays. Specifically, I have shown that reduction of depth magnitude percepts is not a direct consequence of Gestalt grouping. Rather, it is a consequence of object based interpretations of the stimulus which are facilitated only by specific grouping cues that possess spatial support in the form of contours or contour fragments. I have also outlined the ecological importance of such a constraint. This work highlights the importance of investigating the interactions between low level disparity detection and higher level perceptual interpretations of scenes. Future research should focus on investigating the extent of this interaction as well as focus on studying the neural basis for such feedback mechanisms.

## REFERENCES


Attneave, F. (1954). Some informational aspects of visual perception. *Psychological review, 61*(3), 183.

Berry, R. N., Riggs, L. A., & Duncan, C. P. (1950). The relation of vernier and depth discriminations to field brightness. *Journal of Experimental Psychology, 40*(3), 349.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *The Journal of Physiology, 211*(3), 599.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research, 14*(7), 519-527.

Bülthoff, H., Fahle, M., & Wegmann, M. (1991). Perceived depth scales with disparity gradient. *Perception, 20*(2), 145-153.

Bulthoff, I., Bulthoff, H., & Sinha, P. (1998). Top-down influences on stereoscopic depth-perception. *Nature Neuroscience, 1*(3), 254-257.

Burr, D., & Ross, J. (1986). Visual processing of motion. *Trends in Neurosciences, 9*, 304-307.

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45*(7), 891-900.

Christophers, R., & Rogers, B. (1994). *The effect of viewing distance on the perception of complex random-dot stereograms.* Paper presented at the Investigative ophthalmology & visual science.

Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research, 31*(12), 2195-2207.

Coutant, B. E., & Westheimer, G. (1993). Population distribution of stereoscopic ability. *Ophthalmic and Physiological Optics, 13*(1), 3-7.

Deas, L. M., & Wilcox, L. M. (2014). Gestalt grouping via closure degrades suprathreshold depth percepts. *Journal of Vision, 14*(9).

Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research, 33*(7), 981-991.

Elder, J., & Zucker, S. (1994). A measure of closure. *Vision Research, 34*(24), 3361-3369.

Fahle, M., & Westheimer, G. (1988). Local and global factors in disparity detection of rows of points. *Vision Research, 28*(1), 171-178.

Fan, W., Brown, B., & Yap, M. K. (1996). A new stereotest: the double two rod test. *Ophthalmic and Physiological Optics, 16*(3), 196-202.

Fawcett, S. L., & Birch, E. E. (2000). Interobserver test-retest reliability of the Randot preschool stereoacuity test. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 4*(6), 354-358.

Feldman, J., & Singh, M. (2005). Information along contours and object boundaries. *Psychological review, 112*(1), 243.

Foley, J., Applebaum, T., & Richards, W. (1975). Stereopsis with large disparities: discrimination and depth magnitude. *Vision Research, 15*(3), 417-421.

Foley, J., & Richards, W. (1972). Effects of voluntary eye movement and convergence on the binocular appreciation of depth. *Perception & Psychophysics, 11*(6), 423-427.

Gillam, B. (1975). New evidence for "closure" in perception. *Perception & Psychophysics, 17*(5), 521-524.

Glennerster, A., & McKee, S. (1999). Bias and sensitivity of stereo judgements in the presence of a slanted reference plane. *Vision Research, 39*(18), 3057-3069.

Hadad, B.-S., & Kimchi, R. (2008). Time course of grouping of shape by perceptual closure: Effects of spatial proximity and collinearity. *Perception & Psychophysics, 70*(5), 818-827.

Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science, 17*(7), 572-576.

Helmholtz, H. v. (1925). Physiological optics. *Optical Society of America, 3*, 318.

Hodgins, J. (2005). *CMU graphics lab motion capture database*. Retrieved June 4, 2015, from http://mocap.cs.cmu.edu/

Howard, I. P., & Rogers, B. J. (2012). *Perceiving in Depth*: Oxford: Oxford University Press.

Koffa, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt, Brace, & World.

Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences, 90*(16), 7495-7497.

Levinthal, B. R., & Franconeri, S. L. (2011). Common-fate grouping as feature selection. *Psychological Science, 22*(9), 1132–1137.

Lu, H., Tjan, B. S., & Liu, Z. (2006). Shape recognition alters sensitivity in stereoscopic depth discrimination. *Journal of Vision, 6*(1), 7.

Marr, D., & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science, 194*(4262), 283-287.

McKee, S. P. (1983). The spatial requirements for fine stereoacuity. *Vision Research, 23*(2), 191-198.

Mitchison, G. J., & Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research, 24*(9), 1063-1073.

Ogle, K. N. (1952a). Disparity limits of stereopsis. *AMA archives of ophthalmology, 48*(1), 50-60.

Ogle, K. N. (1952b). On the limits of stereoscopic vision. *Journal of Experimental Psychology, 44*(4), 253.

Ogle, K. N. (1953). Precision and validity of stereoscopic depth perception from double images. *Journal of the Optical Society of America, 43*(10), 906-913.

Ogle, K. N., & Weil, M. P. (1958). Stereoscopic vision and the duration of the stimulus. *AMA archives of ophthalmology, 59*(1), 4-17.

Patel, S. S., Bedell, H. E., Tsang, D. K., & Ukwade, M. T. (2009). Relationship between threshold and suprathreshold perception of position and stereoscopic depth. *Journal of the Optical Society of America, 26*(4), 847-861.

Saarinen, J., & Levi, D. M. (1999). The effect of contour closure on shape perception. *Spatial Vision, 12*(2), 227-238.

Saidpour, A., Braunstein, M. L., & Hoffman, D. D. (1994). Interpolation across surface discontinuities in structure from motion Asad Satdpour. *Perception & Psychophysics, 55*(6), 611-622.

Schor, C. M., & Howarth, P. A. (1986). Suprathreshold stereo-depth matches as a function of contrast and spatial frequency. *Perception, 15*(3), 249-258.

Sekuler, A. B. (1990). Motion segregation from speed differences: Evidence for nonlinear processing. *Vision Research, 30*(5), 785-795.

Sekuler, A. B., & Bennett, P. J. (2001). Generalized common fate: Grouping by common luminance changes. *Psychological Science, 12*(6), 437-444.

Shevelev, I. A., Kamenkovich, V. M., & Sharaev, G. A. (2003). The role of lines and corners of geometric figures in recognition performance. *Acta neurobiologiae experimentalis, 63*(4), 361-368.

Sperling, G., Landy, M. S., Dosher, B. A., & Perkins, M. E. (1989). Kinetic depth effect and identification of shape. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 826.

van Boxtel, J. J., & Lu, H. (2013). A biological motion toolbox for reading, displaying, and manipulating motion capture data in research settings. *Journal of Vision, 13*(12), 7-7.

van den Berg, M., Kubovy, M., & Schirillo, J. A. (2011). Grouping by regularity and the perception of illumination. *Vision Research, 51*(12), 1360-1371.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A Century of Gestalt Psychology in Visual Perception: I. Perceptual Grouping and Figure-Ground Organization. *Psychological Bulletin, 138*(6), 1172-1217.

Wallach, H., & Zuckerman, C. (1963). The constancy of stereoscopic depth. *The American journal of psychology, 76*(3), 404-412.

Werner, H. (1937). Dynamics in binocular depth perception. *Psychological Monographs, 49*(2), i.

Wertheimer, M. (1923). Laws of organization in perceptual forms. *A Source Book of Gestalt Psychology*.

Wertheimer, M. (1938). Laws of organization in perceptual forms *A Source Book of Gestalt Psychology* (pp. 71-88). London, England: Kegan Paul, Trench, Trubner & Company.

Wheatstone, C. (1838). Contributions to the physiology of vision.--Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical transactions of the Royal Society of London*, 371-394.

Wist, E. R., Ehrenstein, W. H., & Schraus, M. (1998). A computer-assisted test for the electrophysiological and psychophysical measurement of dynamic visual function based on motion contrast. *Journal of Neuroscience Methods, 80*(1), 41-47.

Woo, G., & Sillanpaa, V. (1979). Absolute stereoscopic thresholds as measured by crossed and uncrossed disparities. *American journal of optometry and physiological optics, 56*(6), 350-355.

Zalevski, A. M., Henning, G. B., & Hill, N. J. (2007). Cue combination and the effect of horizontal disparity and perspective on stereoacuity. *Spatial Vision, 20*(1), 107-138.